



Evaluate Data Center Network Performance

Pilimon, Artur

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Pilimon, A. (2018). *Evaluate Data Center Network Performance*. DTU - Department of Photonics Engineering.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Evaluate Data Center Network Performance

Data center network topology is crucial to latency in the network. Therefore, it is important to know the impact of new topologies before they are deployed.

OVERVIEW

Data centers are the foundation for numerous services that many people today take for granted. Use of these services grows exponentially, causing large organizations to continuously establish new, huge data centers to support the increasing demands.

Data centers contain numerous servers connected through a data center network, which is usually built with layer 2 switches and layer 3 routers. The topology of the data center network is crucial for latency in the data communication to and from the data center and between servers in the data center.

Tests can be conducted to measure latency and other performance parameters for different data center network topologies. It is however important that tests can be repeated and reproduced to have comparable information from the tests.

There are, of course, many topologies that can be used for data center networks. At DTU Fotonik, Department of Photonics Engineering, scientists evaluate data center network topologies with an SDN-based (Software-Defined Networking) control framework measuring network performance – primarily latency. This can be used to plan data center scaling by testing how a new topology will function before changes are made.

Data center network performance can, of course, be tested with Xena Networks solutions. To generate test signals with stateful TCP traffic the Xena Networks testers supporting layer 4-7 - XenaScale and XenaAppliance – are the obvious choice. Testing at lower layers is supported by the XenaBay and XenaCompact test chassis equipped with relevant test modules.

“A case study in how to measure the latency of data center network topologies with an SDN-based control framework and Xena test equipment”

EVALUATE DATA CENTER NETWORK PERFORMANCE

Contents

OVERVIEW	1
INTRODUCTION	3
Data Center Network Testbeds: Performance Evaluation and Testing	5
Data Center Benchmarking	7
DTU Fotonik (Department of Photonics Engineering)	9
Xena Networks Data Center Test Solutions	9
Testing above Layer 3	9
Testing up to Layer 3.....	10
Test Automation	11
CONCLUSION	11

WHITE PAPER

INTRODUCTION

There are thousands of data centers world-wide and they are indispensable for the modern, web oriented society. Data centers are fundamental for numerous cloud based services, which many people today take for granted like streaming, social media networking, e-commerce, on-line banking, Anything-as-a-Service (XaaS) and many more. Large companies such as Google, Microsoft, Apple and Facebook run data centers to provide their offerings to customers and they continue to build new, hyperscale data centers to support exponentially growing demands.

Data centers are built with numerous servers providing a service that is available to end-users (or “clients”). The servers are connected via a data center network, enabling communication between the end users and the servers (“north-south” traffic). The data center network also enable communication between the servers inside the data center (“east-west” traffic).

The data center networks are built with layer 2 switches and layer 3 routers. Early networks could have a hierarchical tree-like topology as depicted in figure 1.

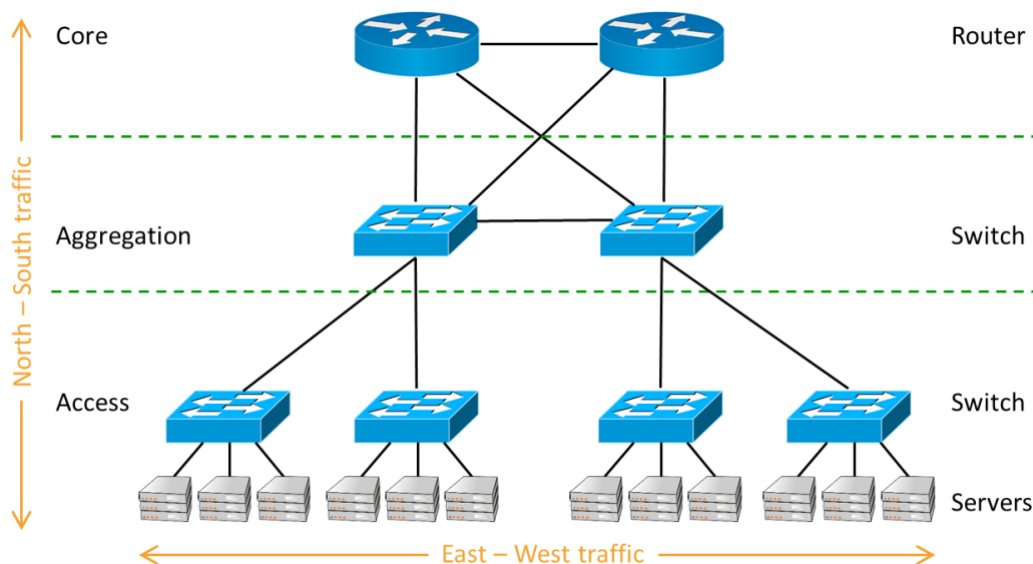


Figure 1: Early data center network topology

With the simple topology in figure 1 there is a risk of congestion in the aggregation layer as several access switches share an aggregation switch. In addition, some of the “east-west” traffic will experience increased latency when the server-to-server traffic must go through several aggregation layer switches. Such issues will increase if more layers are added to support more servers in the data center.

Over the years, the simple data center network topology in figure 1 has developed into other topologies, including the two-layer “leaf-spine” topology, having “leaf” switches forming an Access layer and “spine” switches in the aggregation layer as illustrated in figure 2. With the leaf-spine topology every leaf switch is directly connected to all spine switches in a mesh. Hereby the “east-west” traffic only needs to go through one spine switch, minimizing the latency. The leaf-

spine topology can include spine switches that only handle the “east-west” traffic, reducing the risk for congestion. The leaf-spine topology is useful for data centers with more “east-west” traffic than “north-south” traffic.

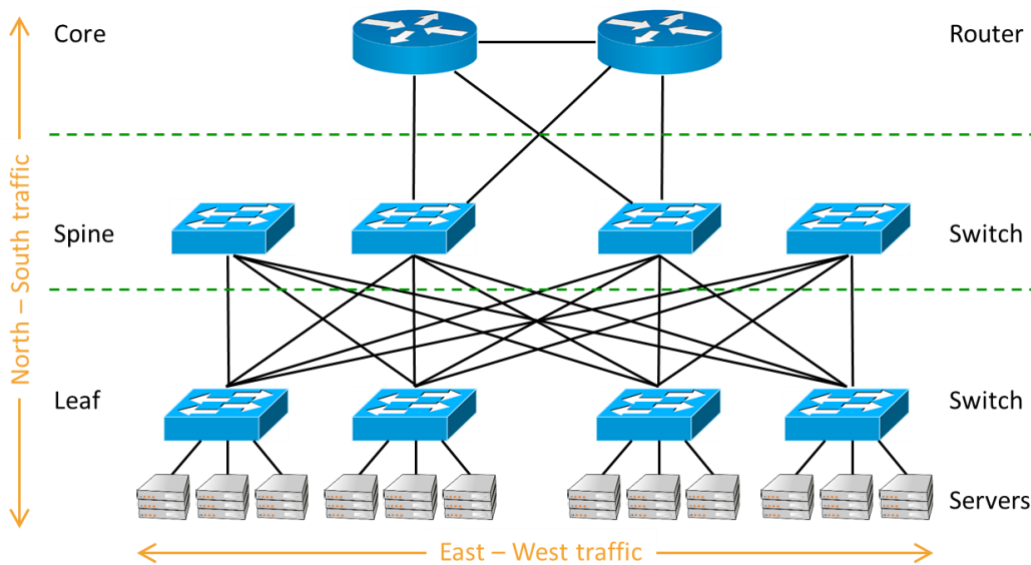


Figure 2: Leaf-spine data center network topology

In figure 1 and figure 2 layer 2 switches are used in the spine/aggregation layer. In some implementations layer 3 routers are used instead. Furthermore, virtualization is widely spread in data centers, meaning that while the logical structure of the data center may be as shown in figures 1 and 2, the actual hardware components may look different.

Software-Defined Networking (SDN) can also be used in data centers. SDN enables very flexible and agile configurations, changing the behavior of network elements by updating their flow tables, which control the traffic forwarding. Hereby the traffic flow through the data center network can dynamically and efficiently be adapted to changing requirements.

Data centers are not isolated entities. In addition to communicating with end-users they can also communicate together through optical Data Center Interconnect (DCI) links over the distance in between them. The increasing need for capacity on these links is driving the development of high speed optical systems.

Latency is an important parameter when designing data center networks. Some applications can be extremely latency sensitive e.g. stock market trading and banking transactions. Therefore, to support latency sensitive applications, data centers network topologies should minimize latency.

Tests can be executed to measure performance and, in particular, latency for different data center network topologies. It is, however, important that tests can be repeated and reproduced in order to have comparable information from the tests.

There are of course more topologies that can be used for data center networks. At DTU Fotonik, Department of Photonics Engineering scientists evaluate data center network topologies with an SDN-based (Software Defined Networking) control framework and measure the performance – primarily latency. This can be used to plan data center scaling by testing the impact of a new topology before changes are made in the data center network.

DATA CENTER NETWORK TESTBEDS: PERFORMANCE EVALUATION AND TESTING

Data centers (DC) and their supporting network infrastructure have become a backbone of the global digital economy, which needs to provide reliable and scalable communication services, while continuously being challenged in terms of the energy efficiency and resource utilization, Quality of Service (QoS) and performance isolation, architectural scalability and cost effectiveness. Hence, it is of paramount importance to conduct timely and comprehensive testing and performance evaluation of the existing and new technologies and protocols, network architectures and traffic engineering (TE) approaches. In addition to analytical modelling and simulations, this can be achieved by applying diverse and innovative research methodologies for testing of real data center network equipment in a realistic communication context (e.g., generating DC-specific traffic profiles, conducting experiments on large-scale DC network testbeds) so that the obtained results could be applicable at scale.

Building a large-scale data center just for experimental research purposes may not be a feasible option, both from the footprint and financial point of view. However, assembling a smaller scale, but sufficiently functional data center testbed, consisting of a subset of real data center network equipment (e.g., electrical and optical switches) with powerful SDN-based control framework and high-performance traffic generators with useful stress-testing capabilities, is a more realistic and flexible approach.

DTU Fotonik – Department of Photonics Engineering at the Technical University of Denmark – is actively using the capabilities of the Layer 2-3 (Xena Bay) and Layer 4-7 (Xena Scale) network testers for their data center research. One of the recent studies carried out was focusing on the experimental evaluation of a direct-connection topology, namely a Hypercube structure, applied as a data center network interconnect, enhanced with optical bypass switching capabilities. High level network connectivity diagrams of two configuration scenarios of a data center testbed are presented in figure 3 (8-node Cube) and figure 4 (16-Hypercube), respectively. All the data center network switches, both optical and electrical, are configured and controlled via an external SDN controller. Nevertheless, one of the most challenging tasks faced in this research activity was to create a functional data center-oriented traffic generation framework to be able to carry out different conformance and performance tests. After multiple different approaches have been tried out, the solution was found by combining the functional capabilities of both Layer 2-3 and Layer 4-7 testers. The reasons are outlined as follows:

1. We were looking for a solution, which could help us achieve two main goals: a) to be able to perform high-speed stress-testing of particular data center network segments and devices by loading these components with large number of traffic flows with sustainable data rates, and

b) to be able to create customized traffic profiles with traffic flow groups of different duration, data volume size as well as configurable network, transport and application layer properties. The former objective was achieved by using Layer 2-3 Xena Bay platform, which also provides great means of collecting accurate per-stream performance statistics (e.g., average, maximum, minimum latency and jitter, packet loss, etc.). The latter requirement was satisfied by using Layer 4-7 Xena Scale tester, which allowed us to mix different groups of stateful (TCP connections) and stateless (UDP flows) traffic flows and configure relevant parameters, such as TCP window sizes, congestion control, segment or flow sizes.

- When conducting network testing at scale, test automation capabilities are becoming critically important, because this results in significant time savings to define and configure various test scenarios as well as process and analyze the gathered results. We used the available CLI-based scripting interface and developed Python scripts to control the tests, gather statistics and visualize the results. Thus, by just changing a set of command-line arguments, a completely different set of tests can be configured automatically. This aspect greatly extends the possibilities of test repeatability and reproducibility of the results.

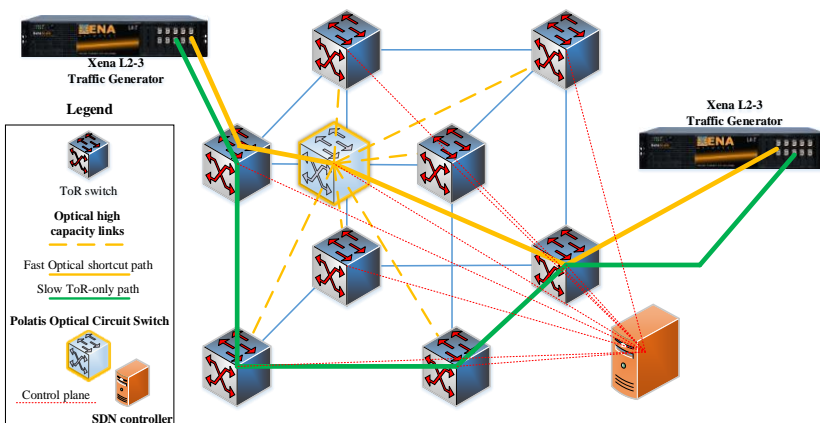


Figure 3: DCN Testbed setup. 8-node Cube with Optical bypass and Xena Bay L2-3 tester

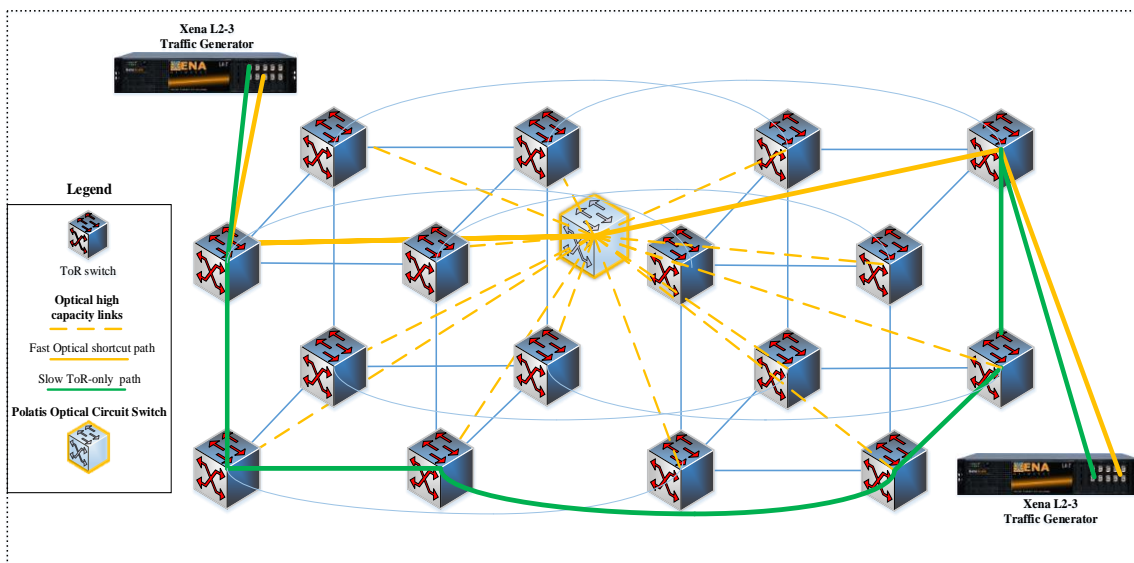


Figure 4: DCN Testbed setup. 16-Hypercube with Optical bypass and Xena Bay L2-3 tester

WHITE PAPER

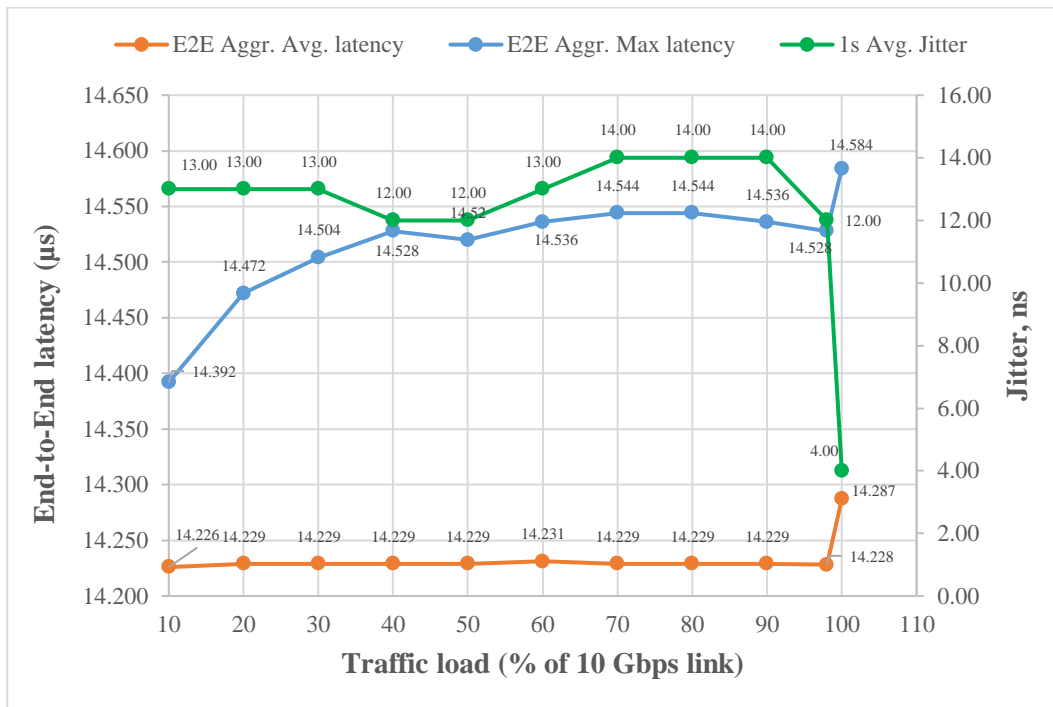


Figure 5: End-to-End latency measurement test results obtained with Xena Bay L2-3 tester. Topology: 8-node Cube, Polatis OCS. Legend: E2E – End-to-End, Aggr. – Aggregate, Avg. – Average, s – second, Max – Maximum

- Another important aspect in favor of using an advanced hardware traffic generator and tester, instead of a software traffic generator, is the resolution (granularity) and accuracy of the measurement results. Considering latency measurements in the context of data center networking, this becomes particularly problematic when using software generators, since the crafted packets are reaching the Network Interface Card (NIC) through the shared kernel space of the underlying operating system (OS), and time-expensive interrupt-based processing is greatly limiting the maximum achievable packet generation rate. Software-level timestamping accuracy is another issue, since simple software generators are limited by the clock granularity of the underlying OS, while hardware counterparts, such as Xena testers, offer more accurate hardware timestamping capabilities and allow tracking statistics at different level of detail, e.g., as it can be seen in figure 5.

DATA CENTER BENCHMARKING

Following the massive deployment of data centers, the Internet Engineering Task Force (IETF) in August 2017 published a couple of documents on data center benchmarking:

- RFC 8238 Data Center Benchmarking Terminology
- RFC 8239 Data Center Benchmarking Methodology

As the titles indicate, RFC 8238 presents new terminology in relation to benchmarking of data center network equipment, while RFC 8239 defines how to perform the benchmarking tests of switches and routers that are used in a data center.

Data center traffic dynamically changes over time. The traffic will in periods predominantly be “north-south” between the servers and a client outside the data center and in other periods be more “east-west” oriented between servers in the data center. Traffic can be a mix of TCP and UDP flows, and can be a result of point-to-multipoint or multipoint-to-multipoint communication patterns, commonly found in data centers. Traffic may be sensitive to latency or throughput and all kinds of traffic can exist simultaneously in a data center network element. Previously IETF has published several documents on network element and network benchmarking, including:

- RFC 2544 Benchmarking Methodology for Network Interconnect Devices
- RFC 2889 Benchmarking Methodology for LAN Switching Devices
- RFC 3918 Methodology for IP Multicast Benchmarking

RFC 8239 have test cases based on the above RFCs. In addition, it includes test cases that better than the above RFCs represent the wide range of traffic conditions that can exist in a data center. RFC 8239 test cases include:

Test	Description
Line-Rate Testing	A "maximum rate" tests for the performance values for throughput, packet drop, latency and jitter. Tests are conducted as a port-pair test and as a full-mesh test
Buffering Testing	Measuring the DUT buffer size under various traffic conditions
Microburst Testing	Identify the maximum amount of packet bursts that a DUT can sustain under various configurations
Head-of-Line Blocking (HOLB)	Examine a DUT's behavior in case of HOLB and measure packet loss caused by HOLB, which occurs when packets are held up by the first packet ahead waiting to be transmitted to a different output port
Incast Stateful and Stateless Traffic	Measure TCP Goodput (retransmissions excluded) and latency under various traffic conditions. The test simulates a mix of stateful (TCP) flows requiring high goodput and stateless (UDP) flows requiring low latency

Table 1: RFC 8239 tests

The RFC 8239 tests require that several (in some cases all) ports of the DUT are connected to a traffic generator, as illustrated in figure 6.

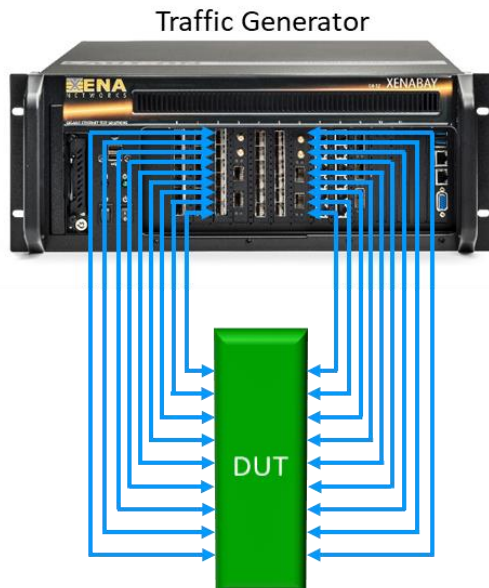


Figure 6: Typical RFC 8239 test setup

DTU FOTONIK (DEPARTMENT OF PHOTONICS ENGINEERING)

At DTU Fotonik, researchers work with multiple aspects of light (photons), in every layer where light can be used and controlled. Approximately 220 researchers work at DTU Fotonik, including around 90 PhD students. Taught programs are B.Sc. Network Technology and IT, M.Sc. in Telecommunications and M.Sc. in Photonics Engineering. In the Network Technologies and Service Platforms group at DTU Fotonik, the research is focused on four main directions, such as data centers (including SDN), fronthaul/backhaul solutions for mobile networks, IoT and core networks. In particular, the group has been leading the recently completed EU FP7 project COSIGN: Combining Optics and SDN In next Generation datacenter Networks (2014 – 2017). The group has sophisticated lab facilities to carry out a variety of datacenter network experiments.

XENA NETWORKS DATA CENTER TEST SOLUTIONS

Data Center Network Performance can of course be tested with Xena Networks test solutions. To generate test signals with stateful TCP traffic the Xena Networks testers supporting layer 4-7 - XenaScale and XenaAppliance – are the obvious choice. Testing at lower layers is supported by the XenaBay and XenaCompact test chassis equipped with relevant test modules.

TESTING ABOVE LAYER 3



Figure 7: The powerful Xena Networks Layer 4-7 testers XenaScale and XenaAppliance

WHITE PAPER

Xena Network's XenaScale and XenaAppliance can be used to generate TCP, HTTP/TCP and UDP traffic streams simultaneously. In addition, both products offer stateful end-to-end testing of network appliances such as switches, firewalls, routers, NAT routers, proxies, load-balancers, bandwidth shapers and more. The platform is also suitable to characterize entire network infrastructure performance for TCP. Top features include:

- Wire-speed stateful TCP traffic generation and analysis with extreme performance
- TLS performance testing with different cipher suites and certificates
- Application emulation with real-world application traffic mixes enabled by XenaAppMix
- Replay captured traffic at scale
- Configuration and tuning of Ethernet, IP and TCP header fields for advanced traffic scenarios
- Stateful TCP connection
- HTTP get/put/head/post
- Extensive live stats and test reports
- 1G – 10G Ethernet interfaces
- 40G Ethernet interfaces (XenaScale)
- High port density – up to 12 x 10 GigE (XenaScale)
- Configurable allocation of processing resources to Ethernet test ports
- Wire-speed traffic capture
- Switched and routed network topologies, NAT support
- Export packet capture to industry standard pcap/Wireshark

TESTING UP TO LAYER 3



Figure 8: The versatile and powerful Xena Networks Layer 2-3 testers XenaBay and XenaCompact

Testing at lower layers is supported by the XenaBay and XenaCompact test chassis equipped with relevant test modules, which can support data rates up to 100 Gbps. Up to 12 test modules can be installed in the XenaBay chassis. Based on Xena's advanced architecture, XenaBay and XenaCompact equipped with relevant test modules are proven solutions for Ethernet testing at layers 2 and 3. Advanced test scenarios can be performed using the free Xena test applications:

- **XenaManager-2G** test software is used to configure and generate streams of Ethernet traffic between Xena test equipment and Devices Under Test (DUTs) and analyze the results
- **Xena2544** offers full support for the 4 test types specified in RFC 2544: Throughput, Latency, Frame loss and Back-to-back frames; Jitter (Frame Delay Variation) is also supported
- **Xena1564** provides full support for both the configuration and performance test types described in Y.1564 for complete validation of Ethernet Service Level Agreements (SLAs) in a single test
- **Xena2889** is an application for benchmarking the performance of Layer 2 LAN switches in accordance with RFC 2889
- **Xena3918** makes it easy to create, edit and execute all test types specified in RFC 3918. RFC 3918 describes tests for measuring and reporting the throughput, forwarding, latency and IGMP group membership characteristics of devices that support IP multicast protocols

TEST AUTOMATION

The Xena Networks L4-7 and L2-3 test solutions have a scripting Command Line Interface (CLI), which is ideal for test automation. The user can create a script, which defines a test sequence that can be repeated as often as required, providing reproducible results.

CONCLUSION

Data centers contain numerous servers providing a service for end-users. The servers are connected via a data center network, which is built with layer 2 switches and layer 3 routers. The topology of the data center network is crucial for latency in the communication between servers and end users and in server-to-server communication.

Many topologies can be used for data center networks. At DTU Fotonik, Department of Photonics Engineering scientists evaluate the performance of different data center network topologies with an SDN-based (Software Defined Networking) control framework, measuring primarily latency.

DTU Fotonik

Department of Photonics Engineering

This can be used to plan data center scaling by testing how a new topology will function before changes are made.

Data Center Network performance can of course be tested with Xena Networks test solutions. To generate test signals with stateful TCP traffic the Xena Networks testers supporting layer 4-7 - XenaScale and XenaAppliance – are the obvious choice. Testing at lower layers is supported by the XenaBay and XenaCompact test chassis equipped with relevant test modules.

